

Title (en)
CALCULATING DEVICE AND CALCULATION METHOD FOR NEURAL NETWORK CALCULATION

Title (de)
BERECHNUNGSVORRICHTUNG UND BERECHNUNGSVERFAHREN ZUR BERECHNUNG MIT NEURONALEM NETZ

Title (fr)
DISPOSITIF DE CALCUL ET PROCÉDÉ DE CALCUL POUR CALCUL DE RÉSEAU NEURONAL

Publication
EP 3561737 A4 20200219 (EN)

Application
EP 17891227 A 20171207

Priority
• CN 201710025196 A 20170113
• CN 2017115038 W 20171207

Abstract (en)
[origin: EP3561737A1] A computing device (300) and a computation method for neural network computation are disclosed. The computing device (300) includes: a first computing unit (310), configured to perform a first operation on an input first matrix M times, to obtain a second matrix; a second computing unit (330), configured to perform a second operation on the input second matrix; and a control unit (340), configured to: control the first computing unit (310) to perform an i -th first operation of the M first operations on the first matrix, to obtain an i -th data element of the second matrix; store the i -th data element of the second matrix into a first storage unit (320); and control, if data elements currently stored in the first storage unit (320) are sufficient for performing one second operation, the second computing unit (330) to perform one second operation. The computing device (300) and the computation method can reduce storage overheads of the computing device (300) for neural network computation.

IPC 8 full level
G06N 3/063 (2006.01)

CPC (source: CN EP US)
G05B 13/027 (2013.01 - US); **G06N 3/063** (2013.01 - CN EP US); **G06N 3/08** (2013.01 - CN US); **G06N 20/10** (2018.12 - US)

Citation (search report)
• [IA] US 2013097212 A1 20130418 - CAI MIKE M [US], et al
• [I] JOHN A ROBINSON: "Efficient Gaussian filtering using Cascaded Prefix Sums", IMAGE PROCESSING (ICIP), 2012 19TH IEEE INTERNATIONAL CONFERENCE ON, IEEE, 30 September 2012 (2012-09-30), pages 117 - 120, XP032333127, ISBN: 978-1-4673-2534-9, DOI: 10.1109/ICIP.2012.6466809
• [A] MAGNUS HALVORSEN: "Hardware Acceleration of Convolutional Neural Networks", 30 June 2015 (2015-06-30), XP055387440, Retrieved from the Internet <URL:https://brage.bibsys.no/xmlui/bitstream/handle/11250/2353511/13656_FULLTEXT.pdf?sequence=1&isAllowed=y> [retrieved on 20170703]
• [A] WANG YU ET AL: "Low power Convolutional Neural Networks on a chip", 2016 IEEE INTERNATIONAL SYMPOSIUM ON CIRCUITS AND SYSTEMS (ISCAS), IEEE, 22 May 2016 (2016-05-22), pages 129 - 132, XP032941496, DOI: 10.1109/ISCAS.2016.7527187
• See references of WO 2018130029A1

Cited by
US10755215B2

Designated contracting state (EPC)
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

Designated extension state (EPC)
BA ME

DOCDB simple family (publication)
EP 3561737 A1 20191030; EP 3561737 A4 20200219; EP 3561737 B1 20230301; CN 108304922 A 20180720; CN 108304922 B 20201215; US 2019340508 A1 20191107; WO 2018130029 A1 20180719

DOCDB simple family (application)
EP 17891227 A 20171207; CN 201710025196 A 20170113; CN 2017115038 W 20171207; US 201916511560 A 20190715